

# ANCOVA

## Kovariančių analizė

# TRUMPAI

**ANCOVA** –kelių grupių vidurkiai lyginami, atsižvelgiant į papildomą informaciją.

**ANCOVA** – tai ANOVA su kovariantėmis.

# PAVYZDŽIAI

- Norima nustatyti, kuris iš dviejų mokymo metodų geresnis, kai, vertinant rezultatus, atsižvelgiama į mokinių motyvaciją ir IQ.
- Tiriama, ar besimokantieji vairuoti, važiuodami dideliais automobiliais, padaro daugiau klaidų nei važiuodami mažais. Papildomai atsižvelgiama į vairuotojo amžių.

# KO TIKIMĖS

- Manome, kad grupių skirtumai išryškės, jei lyginsime to paties intelekto ir vienodos motyvacijos mokinių rezultatus. Nebus taip, kad mažai motyvuoti vienos grupės mokiniai bus lyginami su stipriai motyvuotais kitos grupės mokiniais.
- Tikimės, kad skirtumai išryškės, kai lyginsime maždaug to paties amžiaus vairuotojus.

# PALYGINIMAS SU ANOVA

- ANOVA nulinė hipotezė:

$$H_0 : \mu_1 = \mu_2 = \mu_3.$$

- ANCOVA nulinė hipotezė:

$$H_0 : \mu_{1a} = \mu_{2a} = \mu_{3a};$$

čia  $\mu_{1a}$ ,  $\mu_{2a}$ ,  $\mu_{3a}$  yra grupių vidurkiai,  
„pataisyti“, atsižvelgus į kovariantes.

# PASTABA

Visai įmanoma situacija, kai taikant ANOVA statistiškai reikšmingi skirtumai randami, o taikant ANCOVA – ne. Pavyzdžiui, gali būti, kad pritaikius ANOVA vidutinis vaikų skaičius katalikiškose ir protestantiškose šeimose skirsis. Kita vertus, atsižvelgus į šeimos galvos socialinį ir ekonominį statusą, šie skirtumai dings.

# KINTAMIEJI

- **Priklausomas kintamasis  $Y$**  (t.y. tas kintamasis, kurio vidurkius lyginame) yra intervalinis, normaliai pasiskirstęs.
- **Kovariantė  $X$**  (t.y. Kintamasis, į kurį reikia atsižvelgti) yra intervalinė, normaliai pasiskirsčiusi.
- **Grupavimo** (nepriklausomas, faktorius) **kintamasis** – kategorinis. Jo dėka žinome, kuriai grupei koks respondentas priklauso.

# PAVYZDYS

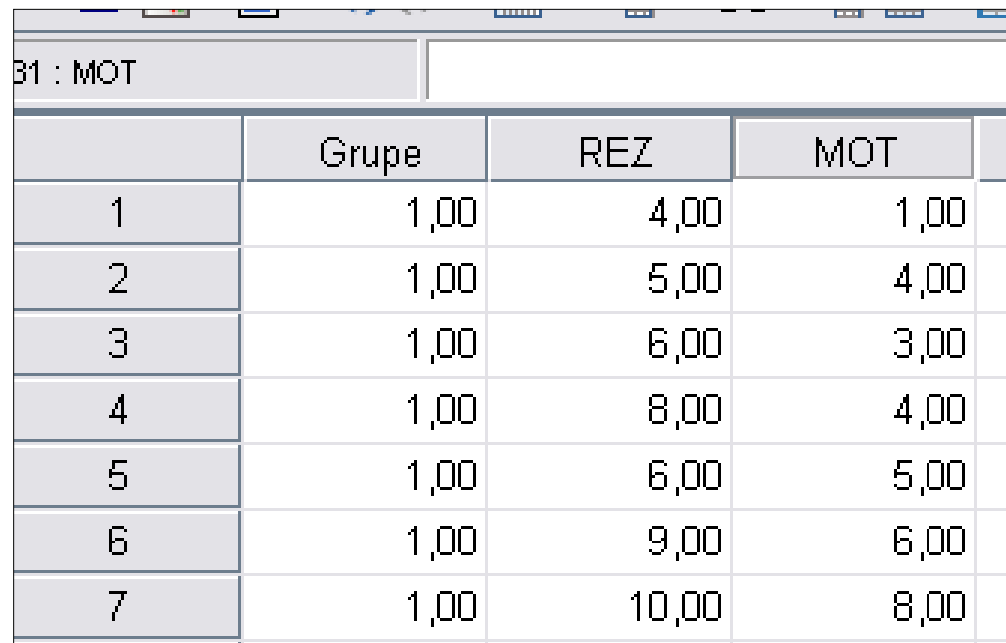
Tiriame, ar statistiškai reikšmingai skiriasi trijų grupių testo rezultatai (**REZ**, intervalinis kintamasis). Atsižvelgsime į respondentų motyvaciją (**MOT**, intervalinis kintamasis). Grupę nurodyto kintamasis **Grupe** (trireikšmis kategorinis kintamasis).

Duomenis galima rasti knygos *V. Čekanaavičius, G. Murauskas, Statistika ir jos taikymai. III*, pirmos dalies šeštajame skyrelyje (1.1.9 pavyzdys).



# SPSS duomenys

Reikia 3 stulpelių. Juose nurodoma kiekvieno respondento grupė (*Grupe*), testo rezultatas (*REZ*) ir motyvacija (*MOT*).

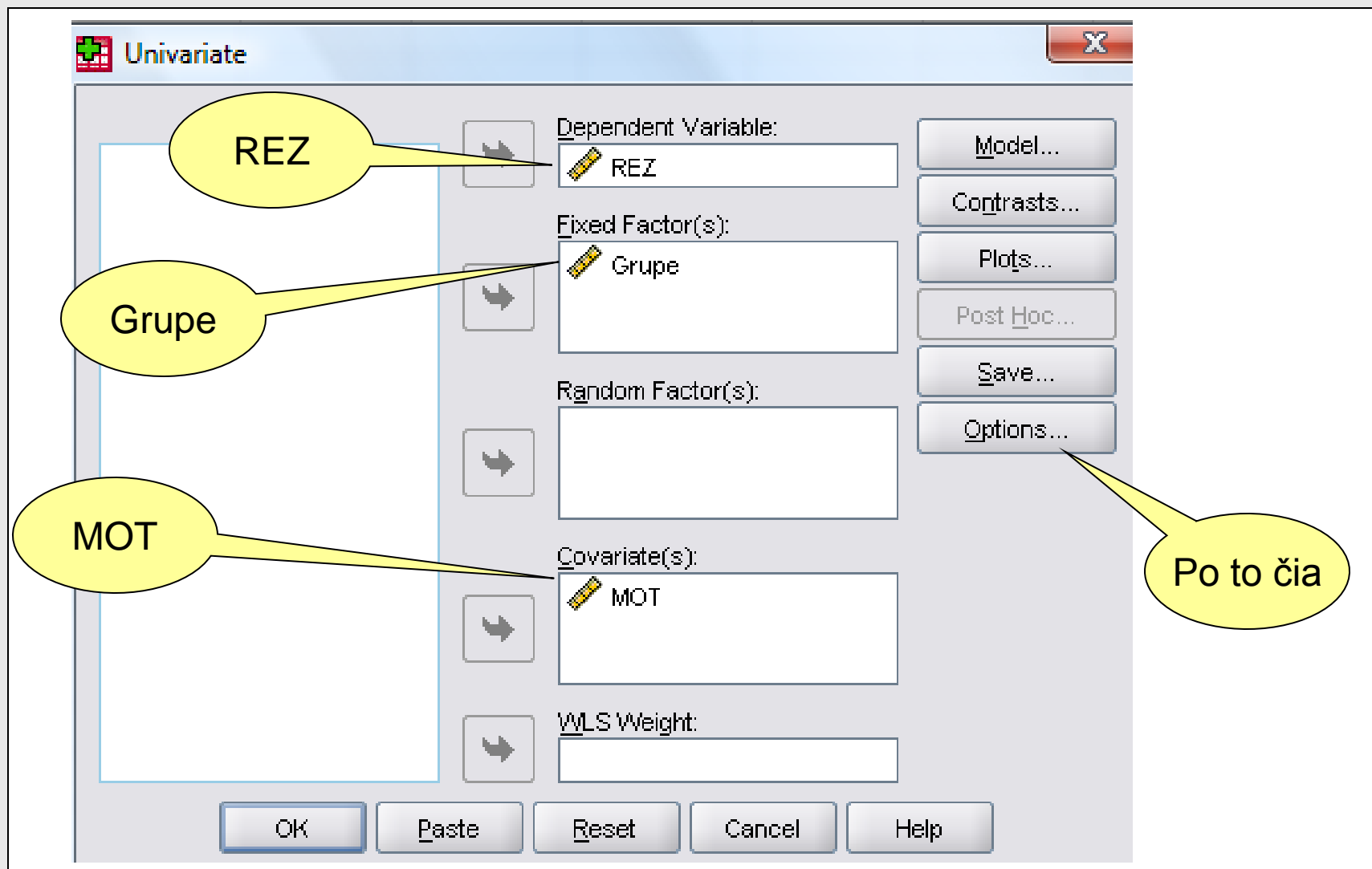


	Grupe	REZ	MOT	
1	1,00	4,00	1,00	
2	1,00	5,00	4,00	
3	1,00	6,00	3,00	
4	1,00	8,00	4,00	
5	1,00	6,00	5,00	
6	1,00	9,00	6,00	
7	1,00	10,00	8,00	

# ANCOVA naudojant SPSS

The image shows the SPSS software interface. The 'Analyze' menu is open, and the 'General Linear Model' option is selected. A secondary menu is open, showing 'GLM GEN Univariate...', 'GLM MULT Multivariate...', and 'GLM REF Repeated Measures...'. A yellow callout bubble points to the 'Univariate...' option with the text 'Čia'. The background shows a data grid with a cell containing the text 'var'.

# ANCOVA naudojant SPSS



# ANCOVA naudojant SPSS

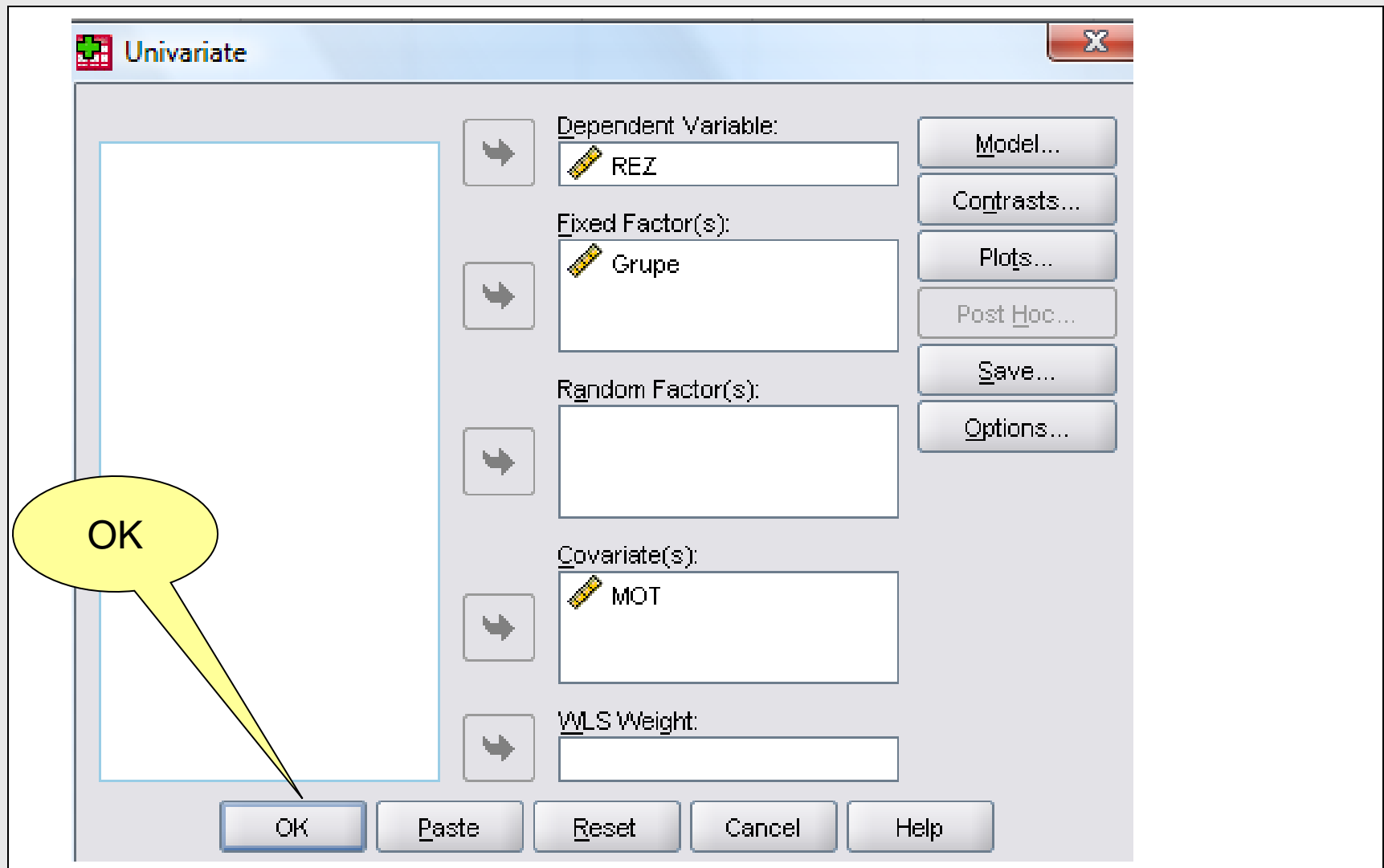
The image shows a screenshot of the 'Univariate: Options' dialog box in SPSS. The dialog is titled 'Univariate: Options' and has a close button (X) in the top right corner. It is divided into several sections:

- Estimated Marginal Means:**
  - Factor(s) and Factor Interactions:** A list containing '(OVERALL)' and 'Grupe'. A blue arrow button is positioned between this list and the 'Display Means for:' list.
  - Display Means for:** A list containing 'Grupe'.
  - Compare main effects
  - Confidence interval adjustment:** A dropdown menu currently set to 'Bonferroni'.
- Display:**
  - Descriptive statistics
  - Estimates of effect size
  - Homogeneity tests
  - Spread vs. level plot

Four yellow callout bubbles with Lithuanian text point to specific elements in the dialog:

- 'Įkeliamė' points to the 'Factor(s) and Factor Interactions:' list.
- 'Uždedame varnelę' points to the 'Compare main effects' checkbox.
- 'Varnelė' points to the 'Descriptive statistics' checkbox.
- 'Pasirenkame' points to the 'Bonferroni' option in the 'Confidence interval adjustment:' dropdown.

# ANCOVA naudojant SPSS



# ANCOVA rezultatai

## Descriptive Statistics

Dependent Variable: REZ

Grupe	Mean	Std. Deviation	N
1,00	8,4000	3,09839	10
2,00	10,6000	3,06232	10
3,00	11,6000	3,13404	10
Total	10,2000	3,28424	30

## Aprašomoji statistika

# ANCOVA rezultatai

## Tests of Between-Subjects Effects

Dependent Variable: REZ

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	293,808 <sup>a</sup>	3	97,936	134,604	,000
Intercept	74,779	1	74,779	102,372	,000
MOT	240,208	1	240,208	328,843	,000
Grupe	54,558	2	27,279	37,344	,000
Error	18,992	26	,730		
Total	3434,000	30			
Corrected Total	312,800	29			

P reikšmė

a. R Squared = ,939 (Adjusted R Squared = ,932)

# ANCOVA rezultatai

Tarkime, kad pasirinkome reikšmingumo lygmenį  $\alpha = 0,05$ . Matome, kad  $p < 0,05$ .

Darome išvada, kad tarp tarp trijų grupių vidutinių testo rezultatų yra statistiškai reikšmingai besiskiriančių.

Liko atsakyti į klausimą, kurių grupių vidutiniai rezultatai skiriasi, t.y. pritaikyti *post hoc* testą.



# ANCOVA rezultatai

Naudojantis Bonferonio kriterijumi, lyginami „pataisyti“ rezultatų vidurkiai, kai atsižvelgiama į motyvacią.

## Estimates

Dependent Variable: REZ

Grupė	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1,00	8,363 <sup>a</sup>	,270	7,808	8,919
2,00	10,673 <sup>a</sup>	,270	10,118	11,229
3,00	11,563 <sup>a</sup>	,270	11,008	12,119

a. Covariates appearing in the model are evaluated at the following values: MOT = 5,7667.

# ANCOVA rezultatai

Ką lyginame

## Pairwise Comparisons

P reikšmės

Dependent Variable: REZ

Su kuo

(I) Grupe	(J) Grupe	Mean Difference (I-J)	Std. Error	Sig. <sup>a</sup>	95% Confidence Interval for Difference <sup>a</sup>	
					Lower Bound	Upper Bound
1,00	2,00	-2,310*	,382	,000	-3,288	-1,332
	3,00	-3,200*	,382	,000	-4,178	-2,222
2,00	1,00	2,310*	,382	,000	1,332	3,288
	3,00	-,890	,382	,084	-1,868	,088
3,00	1,00	3,200*	,382	,000	2,222	4,178
	2,00	,890	,382	,084	-,088	1,868

Vidurkių skirtumai

Based on estimated marginal means

\*. The mean difference is significant at the ,05 level.

a. Adjustment for multiple comparisons: Bonferroni.

# ANCOVA rezultatai

Statistiškai reikšmingai skiriasi tų grupių vidurkiai, kuriuos lyginant  **$p < 0,05$** . Tokių vidurkių skirtumus SPSS automatiškai pažymi žvaigždute.

Gauta, kad pirmoji grupė statistiškai reikšmingai skiriasi nuo likusiųjų, o antroji ir trečioji grupės nesiskiria.

# ANCOVA rezultatų aprašymo pavyzdys

Tyrėme, ar statistiškai reikšmingai skiriasi trijų grupių vidutiniai testo rezultatai. Atsižvelgėme į respondentų motyvaciją. Taikėme kovariančių analizę (ANCOVA) ir Bonferonio kriterijų. Gavome, kad pirmoji grupė statistiškai reikšmingai skiriasi nuo likusiųjų, o antroji ir trečioji grupės nesiskiria.

# PAPILDOMI TYRIMAI

# Palyginimas su ANOVA

Taikant ANCOVA patartina vidurkius palyginti ir be kovariantės. Tam naudojama vienfaktorė dispersinė analizė (ANOVA).

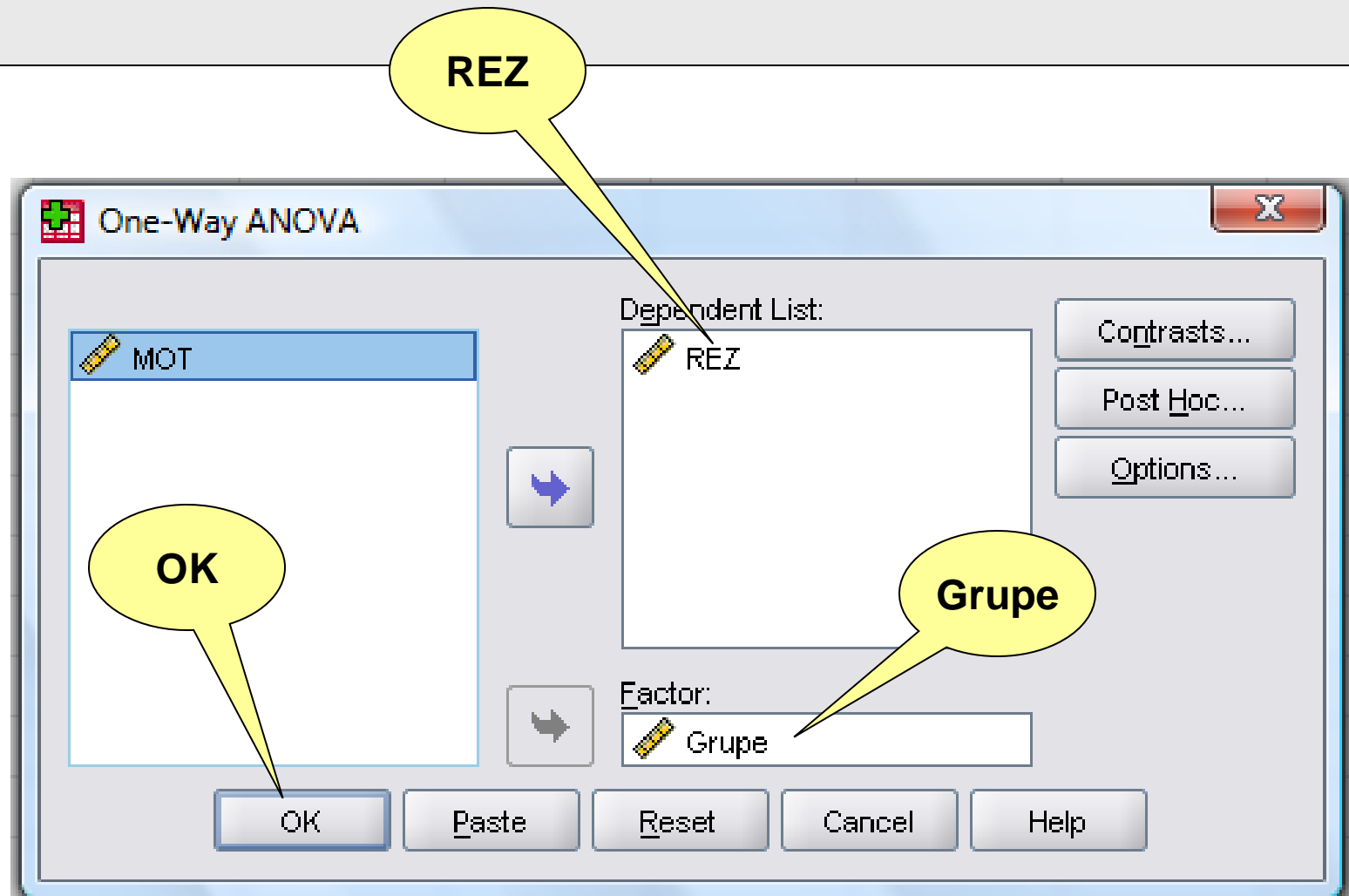
**SPSS** programa ANOVA galima atlikti ne vienu būdu. Naudosime pačią paprasčiausią parinktį.

# Palyginimas su ANOVA

The image shows a software menu structure. The 'Analyze' menu is open, showing options like Reports, Descriptive Statistics, Tables, RFM Analysis, Compare Means, General Linear Model, Generalized Linear Models, Mixed Models, Correlate, and Regression. The 'Compare Means' option is selected, and its sub-menu is open, showing options like Means..., One-Sample T Test..., Independent-Samples T Test..., Paired-Samples T Test..., and One-Way ANOVA... The 'One-Way ANOVA...' option is highlighted in blue. A yellow callout bubble points to this option with the text 'Čia'.

var	var	var
-----	-----	-----

# Palyginimas su ANOVA





# Palyginimas su ANOVA

Vidurkių skirtumai statistiškai nereikšmingi.  
Primename, kad, atsižvelgus į motyvaciją,  
statistiškai reikšmingų skirtumų atsirado.

## ANOVA

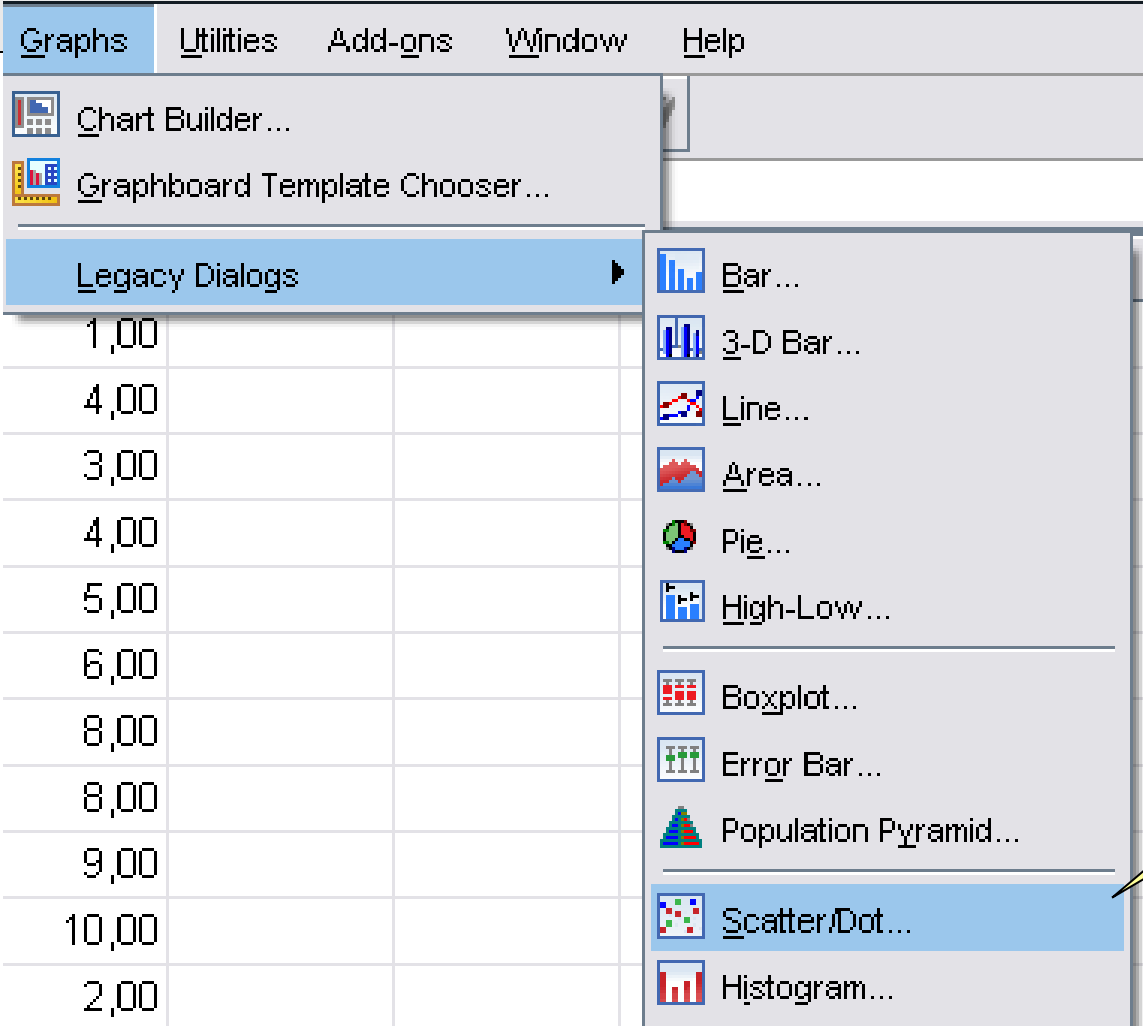
REZ

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	53,600	2	26,800	2,792	,079
Within Groups	259,200	27	9,600		
Total	312,800	29			

# Reikšmių išsibarstymo grafikas

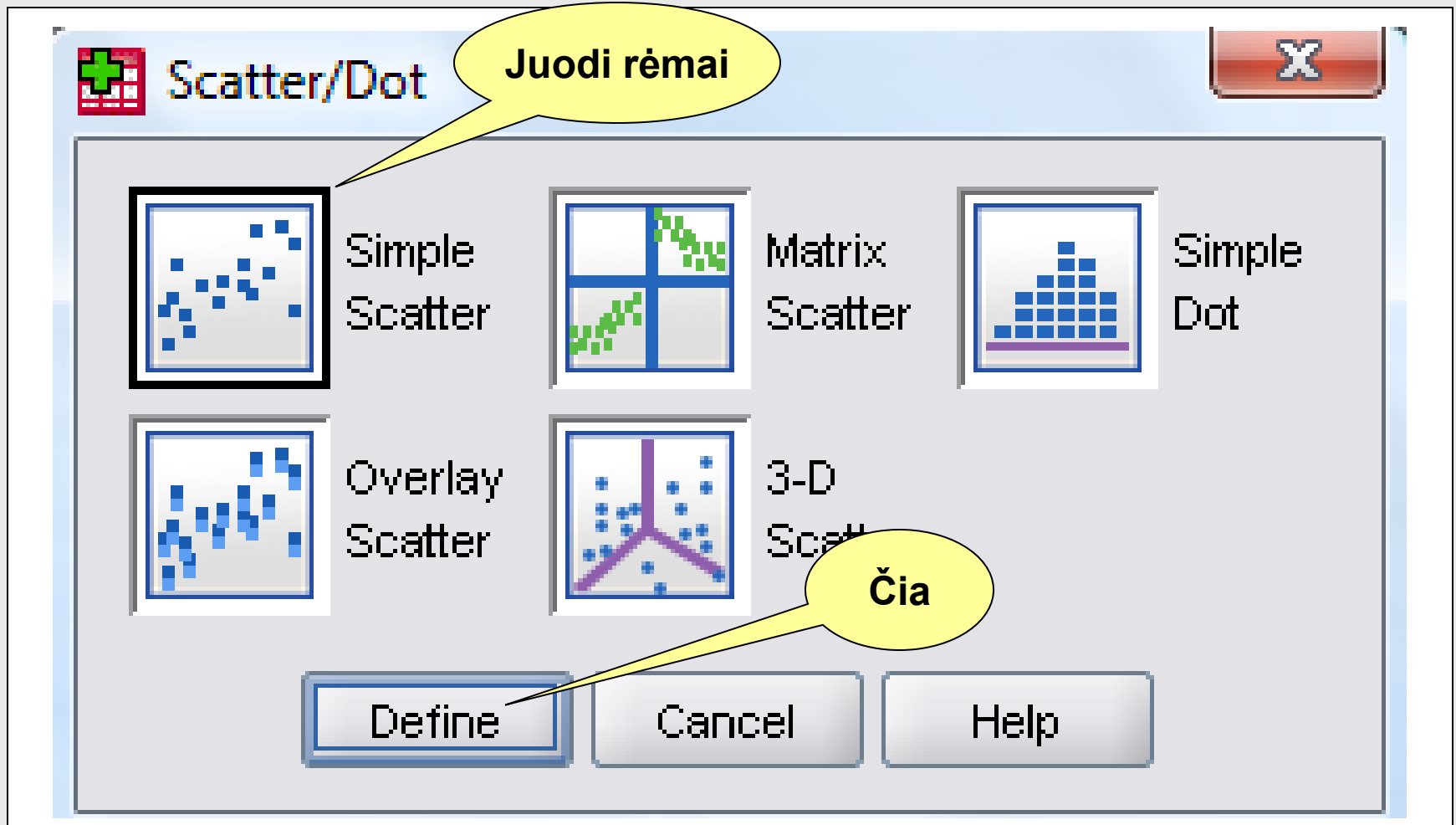
Norint nustatyti testo rezultatų ir motyvacijos priklausomybę, verta nubraižyti abiejų kintamųjų grafiką. Parodysime, kaip tai galima padaryti naudojantis SPSS 17 programa.

# Reikšmių išsibarstymo grafikas

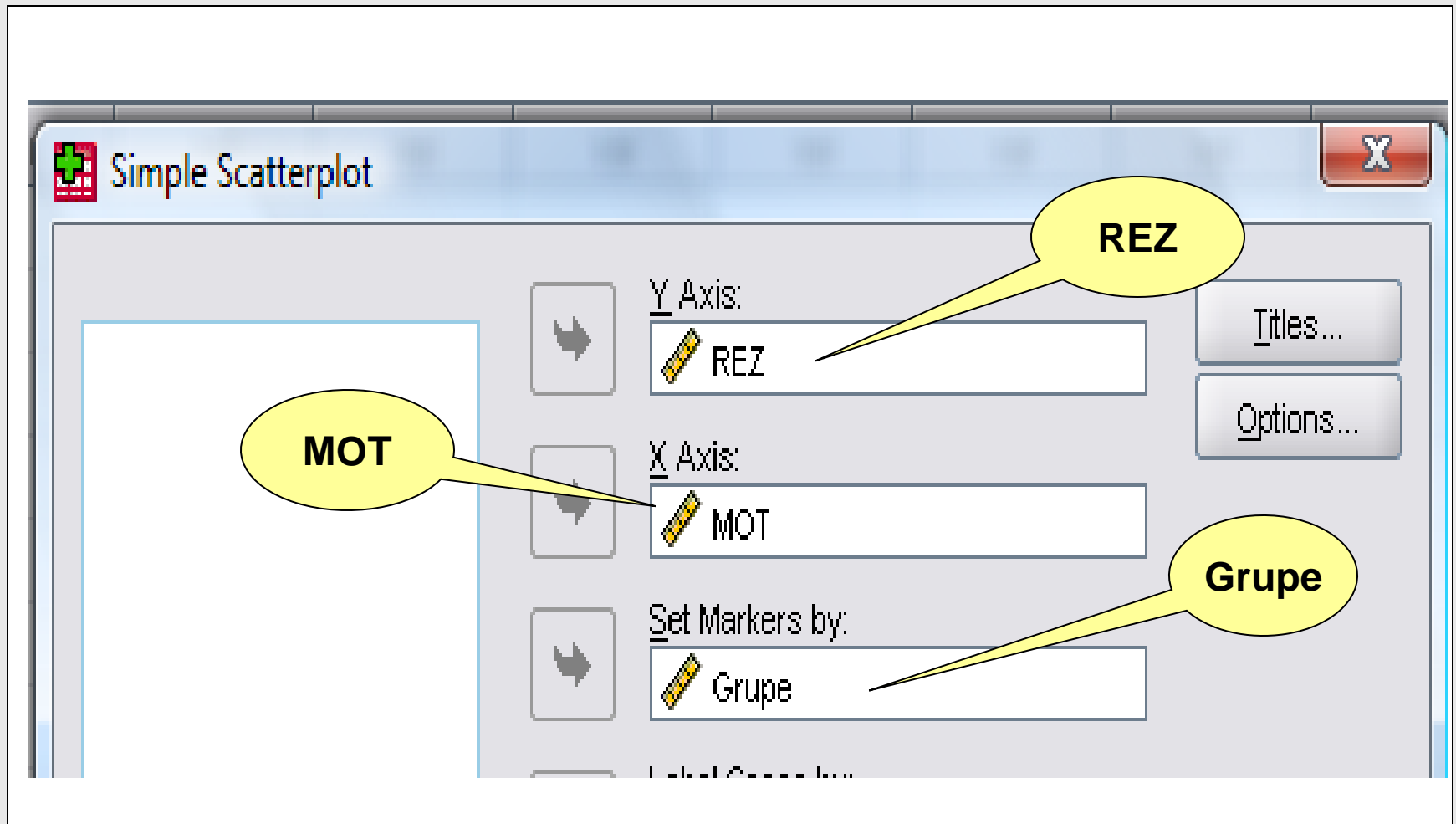


The screenshot shows a software interface with a menu bar containing 'Graphs', 'Utilities', 'Add-ons', 'Window', and 'Help'. The 'Graphs' menu is open, displaying several options: 'Chart Builder...', 'Graphboard Template Chooser...', and 'Legacy Dialogs'. The 'Legacy Dialogs' submenu is also open, listing various chart types: 'Bar...', '3-D Bar...', 'Line...', 'Area...', 'Pie...', 'High-Low...', 'Boxplot...', 'Error Bar...', 'Population Pyramid...', 'Scatter/Dot...', and 'Histogram...'. The 'Scatter/Dot...' option is highlighted in blue. A yellow speech bubble with the word 'Čia' (Here) points to this option. In the background, a grid is visible with numerical values: 1,00, 4,00, 3,00, 4,00, 5,00, 6,00, 8,00, 8,00, 9,00, 10,00, and 2,00.

# Reikšmių išsibarstymo grafikas

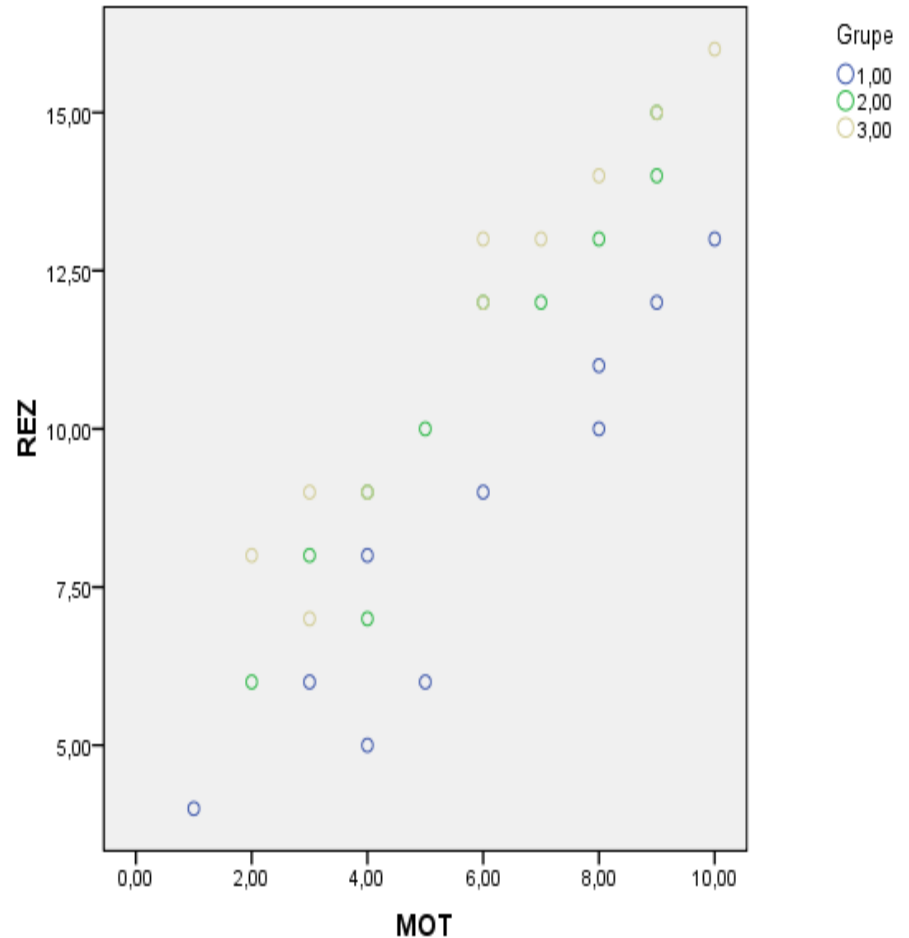


# Reikšmių išsibarstymo grafikas

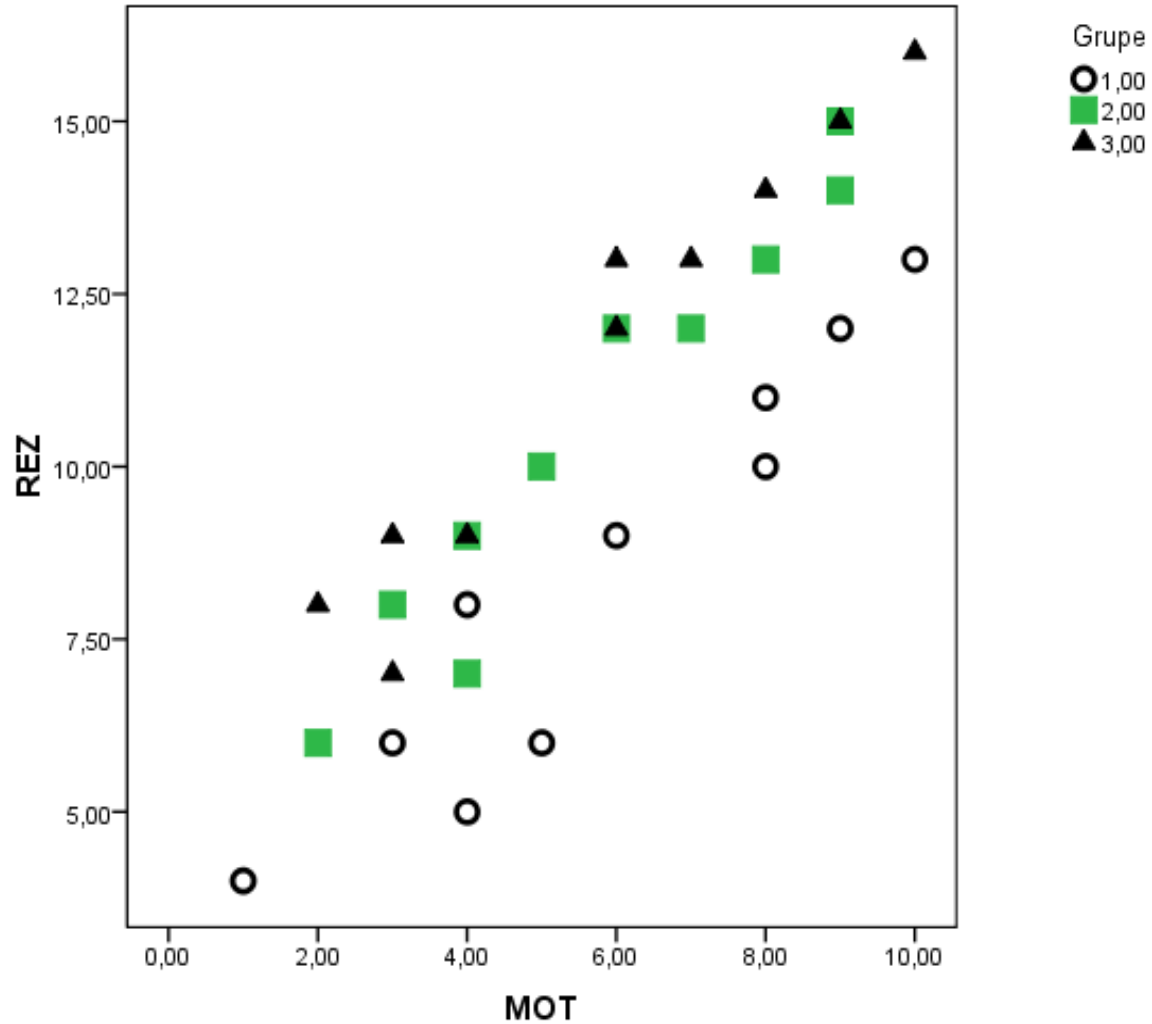


# Reikšmių išsibarstymo grafikas

Grafikas – ne itin vaizdus, tačiau jį galima redaguoti.



# Suredaguotas grafikas



# Į ką grafike atkreipti dėmesį

1. Visose grupėse matyti ta pati tendencija – kuo didesnė motyvacija, tuo geresni rezultatai.
2. Matome, kad skirtingų grupių rezultatų „debesėliai beveik lygiagretūs“. Vadinasi grupės ir motyvacijos sąveikos (interakcijos) nėra. Visose grupėse motyvacijos poveikis maždaug vienodas.
- 3) Kuo labiau debesėliai atsiskiria, tuo tikėtiniau, taikant ANCOVA, rasti statistiškai reikšmingų vidurkių skirtumų.



# Kintamųjų sąveika

Statistiškai reikšminga motyvacijos ir grupės sąveika parodo, kad skirtingose grupėse motyvacija skirtingai veikia rezultatus.

Klasikinė ANCOVA daroma, kai nėra grupės ir kovariantės sąveikos (interakcijos).

Parodysime, kaip naudojant SPSS galima patikrinti, ar sąveikos nėra.

Pradžią nesiskiria nuo įprastinės ANCOVA.

# ANCOVA su kintamųjų sąveika

The image shows a screenshot of the SPSS software interface. The 'Analyze' menu is open, and the 'General Linear Model' option is selected. A secondary menu is open, showing 'Univariate...', 'Multivariate...', and 'Repeated Measures...'. A yellow speech bubble points to the 'Univariate...' option with the text 'Čia' (Here).

The menu structure is as follows:

- Analyze
  - Reports
  - Descriptive Statistics
  - Tables
  - RFM Analysis
  - Compare Means
  - General Linear Model**
    - GLM GEN Univariate...**
    - GLM MULT Multivariate...
    - GLM REF Repeated Measures...
  - Generalized Linear Models
  - Mixed Models

The background shows a data grid with a cell containing the text 'var'.

# ANCOVA su kintamųjų sąveika

The image shows the 'Univariate' dialog box in SPSS. The 'Dependent Variable:' field contains 'REZ'. The 'Fixed Factor(s):' field contains 'Grupe'. The 'Covariate(s):' field contains 'MOT'. The 'WLS Weight:' field is empty. On the right side, there are buttons for 'Model...', 'Contrasts...', 'Plots...', 'Post Hoc...', 'Save...', and 'Options...'. At the bottom, there are buttons for 'OK', 'Paste', 'Reset', 'Cancel', and 'Help'. Four yellow callout bubbles are present: one pointing to 'REZ' in the dependent variable field, one pointing to 'Grupe' in the fixed factor field, one pointing to 'MOT' in the covariate field, and one pointing to the 'Options...' button.

REZ

Grupe

MOT

Po to čia

# ANCOVA su kintamųjų sąveika

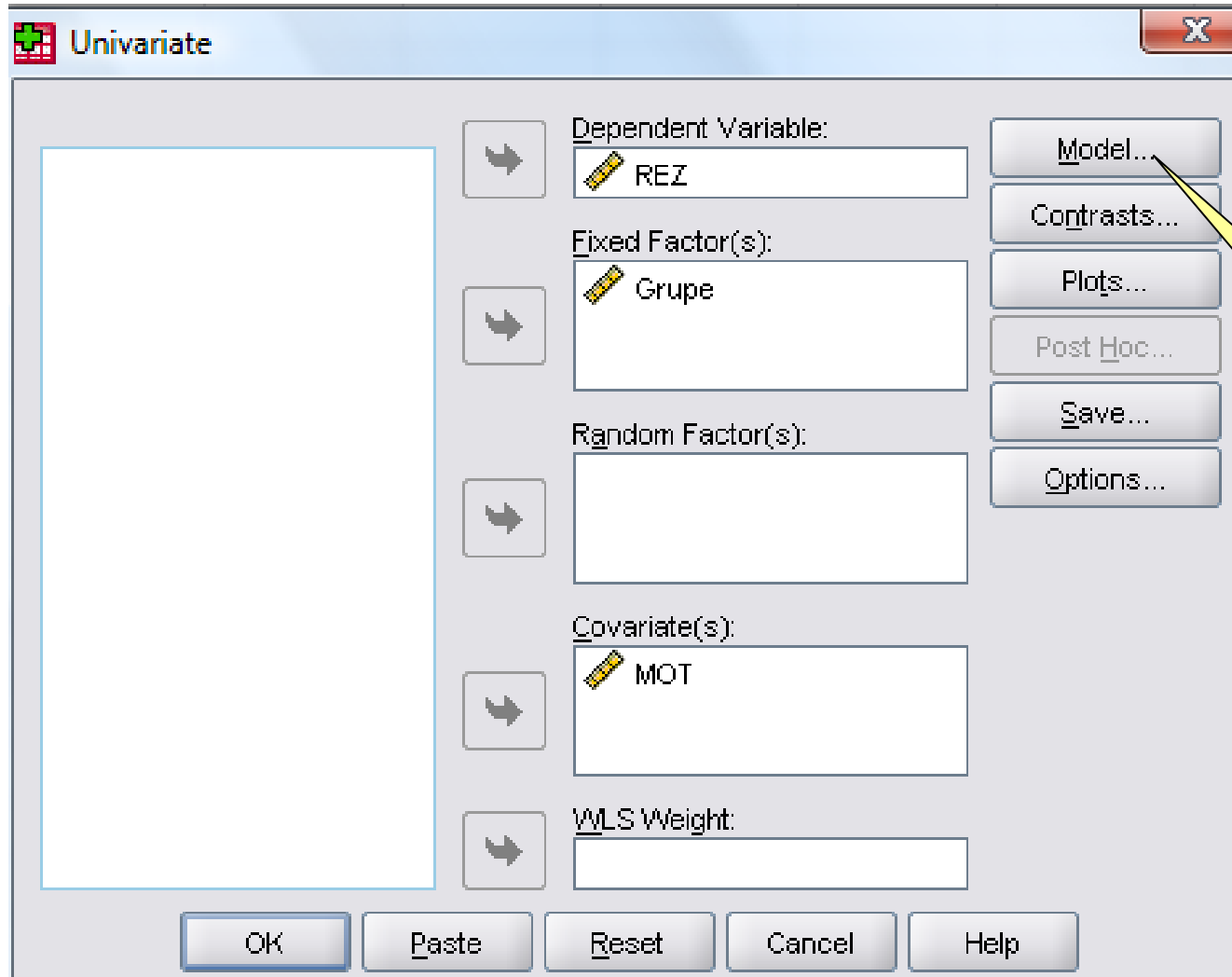
The image shows a screenshot of the 'Univariate: Options' dialog box in SPSS. The dialog is titled 'Univariate: Options' and has a close button (X) in the top right corner. It is divided into several sections:

- Estimated Marginal Means:**
  - Factor(s) and Factor Interactions:** A list box containing '(OVERALL)' and 'Grupe'. A blue arrow button is positioned between this list and the 'Display Means for:' list.
  - Display Means for:** A list box containing 'Grupe'. A yellow callout bubble labeled 'Įkeliamo' (Loading) points to this list.
  - Compare main effects
  - Confidence interval adjustment:** A dropdown menu set to 'Bonferroni'. A yellow callout bubble labeled 'Pasirenkame' (We choose) points to this dropdown.
- Display:**
  - Descriptive statistics
  - Homogeneity tests
  - Estimates of effect size
  - Spread vs. level plot

Other annotations include:

- A yellow callout bubble labeled 'Uždedame varnelę' (We add a checkmark) pointing to the 'Compare main effects' checkbox.
- A yellow callout bubble labeled 'Varnelė' (Checkmark) pointing to the 'Descriptive statistics' checkbox.

# ANCOVA su kintamųjų sąveika



# ANCOVA su kintamųjų sąveika

The screenshot shows the 'Univariate: Model' dialog box in SPSS. The 'Specify Model' section has the 'Custom' radio button selected, with a yellow callout bubble labeled 'Taškas' pointing to it. The 'Factors & Covariates' list on the left contains 'Grupe' and 'MOT'. The 'Build Term(s)' section has a dropdown menu set to 'Interaction' and a blue arrow button. The 'Model:' list on the right contains 'Grupe', 'MOT', and 'Grupe\*MOT', with 'Grupe\*MOT' highlighted in blue and a yellow callout bubble labeled 'Sąveika' pointing to it. Another yellow callout bubble labeled 'Įkeliamė' points to the 'Grupe' and 'MOT' terms in the 'Model:' list.

Univariate: Model

Specify Model

Full factorial  Custom

Factors & Covariates:

- Grupe
- MOT

Build Term(s)

Type:

Interaction

Model:

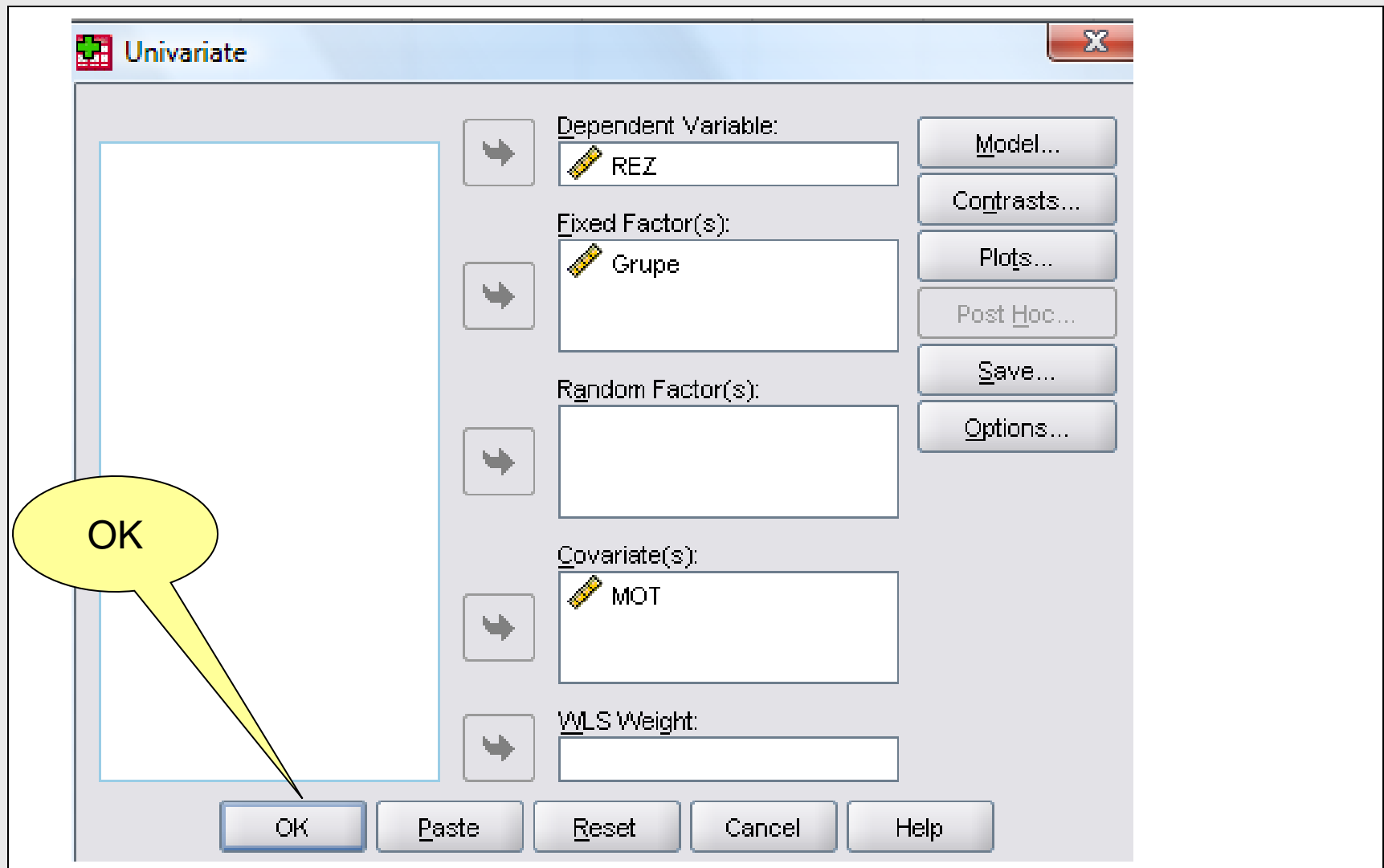
- Grupe
- MOT
- Grupe\*MOT

Taškas

Įkeliamė

Sąveika

# ANCOVA su kintamųjų sąveika



# ANCOVA su kintamųjų sąveika

## Tests of Between-Subjects Effects

Dependent Variable: REZ

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	294,814 <sup>a</sup>	5	58,963	78,680	,000
Intercept	72,200	1	72,200	96,344	,000
Grupe	6,102	2	3,051	4,071	,030
MOT	240,283	1	240,283	320,632	,000
Grupe * MOT	1,006	2	,503	,671	,520
Error	17,986	24	,749		
Total	3434,000	30			
Corrected Total	312,800	29			

a. R Squared = ,943 (Adjusted R Squared = ,931)

Sąveikos  
nėra.



## Ką daryti, jeigu sąveika yra

Tuo atveju, interpretuojant rezultatus, reikia kalbėti ne apie ANCOVA, bet apie tiesinį Gauso modelį (LNM). Reikėtų užrašyti modelį

$$\text{REZ} = \beta_0 + \beta_1 \text{MOT} + \beta_2 \text{Grupe} + \beta_3 \text{Grupe} \cdot \text{MOT} + e$$

ir jį komentuoti (priklauso nuo motyvacijos, grupės ir motyvacijos sąveikos su grupe ir pan.).

# Baigiamosios pastabos

Tyrėme vienfaktorę ANCOVA, t.y. modelį su viena kovariante. Kovariančių gali būti ir daugiau. Modelio analizė nuo to nesikeičia. Vis dėlto, kuo mažiau kovariančių, tuo modelį lengviau interpretuoti.